Award Number: W81XWH-12-1-0391

TITLE: Meta-Analytical Online Repository of Gene Expression Profiles of MDS Stem Cells

PRINCIPAL INVESTIGATOR: Dr. Amit Verma, M.D.

CONTRACTING ORGANIZATION: Albert Einstein College of Medicine
Bronx, NY 10461

REPORT DATE: October 2014

TYPE OF REPORT: ANNUAL

PREPARED FOR: U.S. Army Medical Research and Material Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2014	Annual	30 Sep 2013 - 29 Sep 2014

4. TITLE AND SUBTITLE
Meta-Analytical Online Repository of Gene Expression Profiles of MDS Stem Cells

5a. CONTRACT NUMBER
W81XWH-12-1-0391

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Amit Verma

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

E-Mail: amit.verma@einstein.yu.edu

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Albert Einstein College of Medicine
Cancer Center
1300 Morris Park Avenue
Bronx, NY 10461-1602

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

14. ABSTRACT

We proposed to create an repository of gene expression profiles of CD34+ stem cells from patients with myelodysplastic syndromes (MDS) and healthy controls. We have devised a novel meta-analytical approach to integrate and normalize gene expression studies generated in different labs on different platforms and have shown it to be feasible and biologically valid in numerous publications. We have now successfully integrated data from 183 MDS samples in the database. This dataset has been used in numerous studies. These include a screen for expression of all TGF-β related genes and has led to the discovery that negative regulator, SMAD7, was significantly underexpressed in MDS stem cells. A recent study using this database has led to the discovery of STAT3 as a therapeutic target in MDS stem cells. We are now using this database to screen for chemokine genes in MDS and have identified the chemokine IL8 as a promising target that is upregulated in MDS. We have now added information about MDS subtypes, blood counts, IPSS scores, patient demographics to this database. We now propose to add information about mutations in these samples. This resource will continually be updated with newer data on an ongoing basis.

15 SUBJECT TERMS
MDS, Stem Cells, Gene Expression

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			
U	U	U	UU		19b. TELEPHONE NUMBER (include area code)

Table of Contents

Introduction:

MDS is a heterogeneous group of diseases characterized by bone marrow failure that leads to cytopenias. Newer therapeutic developments are impeded by limited insights into disease pathophysiology. Lack of cell lines and valid mouse models underscore the importance of research based on primary patient samples. These samples are hard to obtain in sufficient numbers. Our proposal will create a gene expression database that will include large numbers of all MDS subtypes. Moreover, this database will be generated by compiling gene expression profiles from CD34+ purified stem cells, thus ensuring that these profiles are not diluted due to heterogeneity of whole bone marrow samples. In addition to allowing researchers to study the expression patterns of selected genes, this database will also allow researchers to identify subtypes of MDS patients that will potentially benefit from novel therapies that target specific genes or genetic pathways.

Body:

### Scientific accomplishments in previous funding cycle:

We had previously shown that meta-analysis of microarray studies demonstrates feasibility of inter-platform data integration and reveals novel hematopoietic stem cell signatures. We tested the feasibility of conducting a meta-analysis of GE studies by using publically available data from studies that used normal bone marrow-derived hematopoietic progenitors. Data was integrated using both RefSeq and UniGene identifiers and normalized. We observed that in spite of variability introduced by experimental conditions and different microarray platforms, our meta-analytical approach can distinguish biologically distinct normal tissues by clustering them based on their cell of origin(1).

After demonstrating the feasibility of our meta-analytical approach, we wanted to construct a database of MDS stem cells and normal controls. We have now integrated data from 183 MDS CD34+ samples and 17 healthy controls.   The data was integrated using unigene IDs, normalized and shown to be valid for further analysis (2). We have subsequently used this database in numerous studies that have led to important insights into the pathophysiology of MDS [1-4]. Studies in the previous funding cycle evaluated TGF-beta regulated genes in MDS and established SMAD7 reduction as a key intracellular event that leads to myelosuppressive TGF-β signaling and ineffective hematopoiesis in MDS [1, 4]. In another previous study, we determined that DOCK4, a GTPase exchange factor is underexpressed in MDS [3]

### Scientific accomplishments in last years funding cycle:

Identification of STAT3 overexpression in MDS: We have now been able to use the database to screen for transcription factors that are overexpressed in MDS. We identified STAT3 as a promising target, that is overexpressed in large number of MDS stem cells (Fig 1). This data was used in a recent paper published in Blood [2].

We have also incorporated data for patient survival in the database. Analysis of STAT3 and survival shows that high STAT3 is marker of adverse prognosis (Fig 1, right panel). This data is presently unpublished.
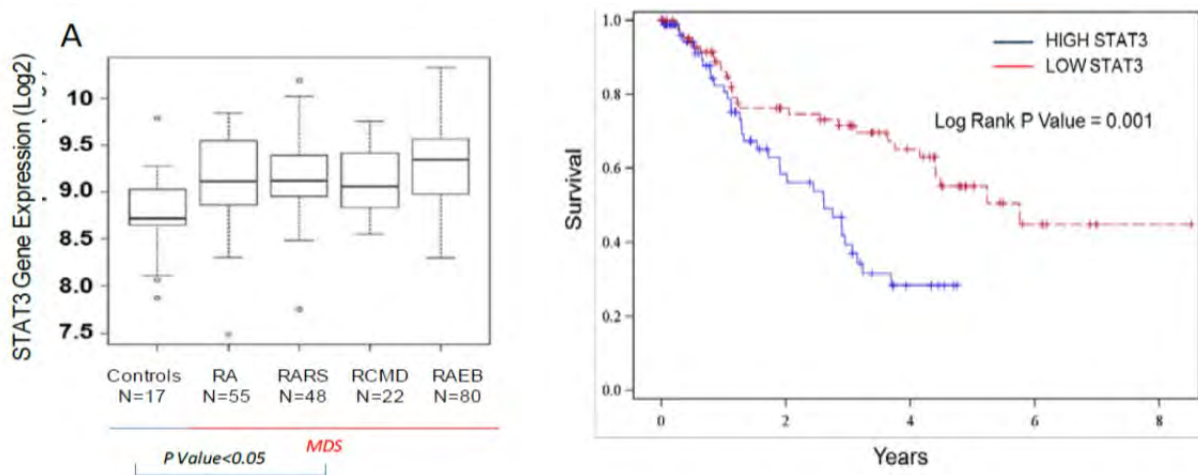
Figure 1: **STAT3 is overexpressed in MDS stem cells and is marker of worse prognosis:** STAT3 expression is significantly increased in 183 MDS CD34+ samples when compared to 17 healthy controls. Box plots shows expression in histological subtypes of MDS (TTest with Benjamin Hochberg correction) (A). Survival of 183 MDS patients was correlated with STAT3 expression in CD34+ cells. Patients with higher STAT3 levels (> median) had a median survival of 2.6 years compared to 5.8 years for group with lower STAT3 (Log Rank P Value<0.01). (Right sided panel)

Identification of CXCR2 overexpression in MDS: We have been able to use the database to screen for chemokines that are overexpressed in MDS. We identified the CXCR2 receptor as a promising target, that is overexpressed in large number of MDS stem cells (Fig 2). This data was presented at the American society of hematology meeting in 2013. We are now completing experiments and will submit this manuscript very soon.
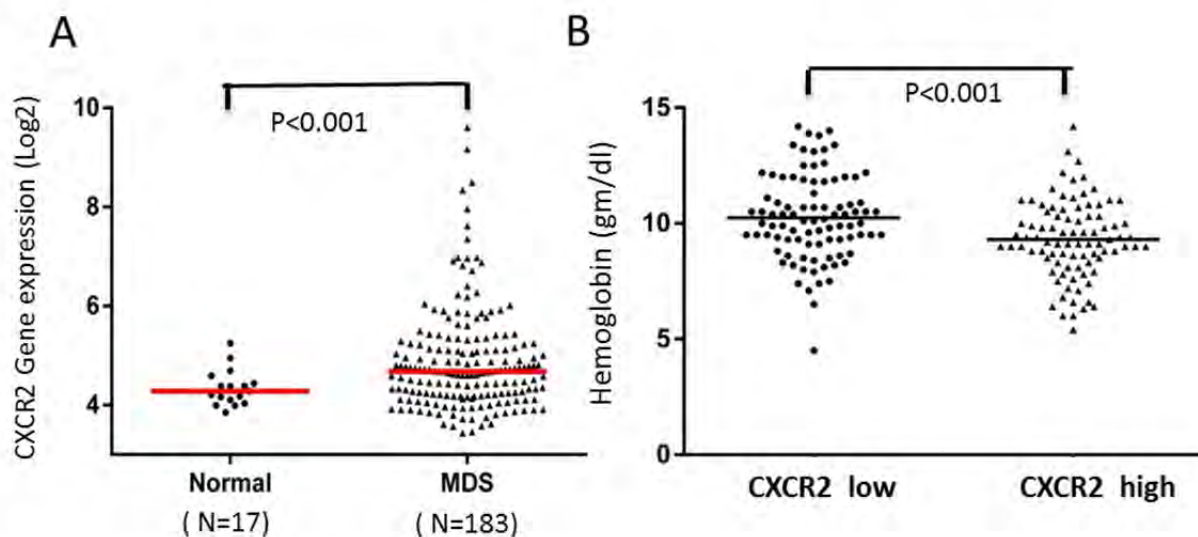
Fig 2: CXCR2 is expressed in MDS samples : Expression of Chemokine ligands and receptors was evaluated in gene expression data from 183 MDS CD34+ cells and 17 healthy control CD34+s. Significantly higher expression of CXCR2 in MDS (Ttest, P Value <0.001, FDR<5%) was observed (A). MDS patients with high CXCR2 expression (>median) have worse hemoglobin (B)(Ttest P Value<0.001)

Mutational profiling of MDS cases in the database: Mutational profiling of the cases is now partially accomplished and preliminary results were published recently [5].   We are now in the process of mutationally profiling all cases and incorporating them in the database.

Proposed work for the last year of funding:

1. Completion of mutational profiling of all cases and incorporation of this data in the database

2. Incorporating patient survival into the database and sharing it with other investigators
3. Incorporating red cell transfusion requirements in the database to provide more clinical correlation for investigators
4. Adding more MDS samples in the database.

Key Research Accomplishments:

We have shown the feasibility of constructing a meta-analytical database of MDS stem cell samples and controls, and have shown that this database can be used in numerous basic as well as translation studies in MDS.

Reportable Outcomes:

Manuscripts:

Hosono N, Makishima H, Jerez A, Yoshida K, Przychodzen B, McMahon S, Shiraishi Y, Chiba K, Tanaka H, Miyano S, Sanada M, Gómez-Seguí I, **Verma AK**, McDevitt MA, Sekeres, Ogawa S, Maciejewski JP. Recurrent genetic defects on chromosome 7q in myeloid neoplasms. *Leukemia* 2014 Jan 16. PMID: 24429498

Bhagat T, Zhou L, Sokol L, Caceres G, Gundabolu K, Gordon S, Mantzaris I, Gligich O, Yu Y, Bhattacharyya S, Jing X, Polineni R, Tamari R, Bhatia K, Pellagatti A , Boultwood J, Kambhampati S, Steidl U, Stein C, Ju W, Liu G, Kenny P, List A, Bitzer M, **Verma A.** miR-21 mediates hematopoietic suppression in MDS by activating TGF- β signaling *Blood* 2013, Apr 11;121(15):2875-81. PMID: 23390194

Will B, Zhou L, Vogler TO, Ben-Neriah S, Schinke C, Tamari R, et al. Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations. *Blood*. 2012; **120**(10): 2076-86.

Gerstung M, Pellagatti A, Malcovati L, Giagounidis A, Porta MG, Jadersten M, et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nature communications*. 2015; 6: 5901.

Conclusion:

We have shown the feasibility of constructing a meta-analytical database of MDS stem cell samples and controls, and have shown that this database can be used in basic as well as translation studies in MDS. We will now incorporate mutational and more clinical information to this database in the new phase of the funding period.

References:

1. Bhagat TD, Zhou L, Sokol L, Kessel R, Caceres G, Gundabolu K, et al. miR-21 mediates hematopoietic suppression in MDS by activating TGF-beta signaling. Blood. 2013; **121**(15): 2875-81.
2. Will B, Zhou L, Vogler TO, Ben-Neriah S, Schinke C, Tamari R, et al. Stem and progenitor

cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations. Blood. 2012; **120**(10): 2076-86.

3.   Zhou L, Opalinska J, Sohal D, Yu Y, Mo Y, Bhagat T, et al. Aberrant epigenetic and genetic marks are seen in myelodysplastic leukocytes and reveal Dock4 as a candidate pathogenic gene on chromosome 7q. The Journal of biological chemistry. 2011; **286**(28): 25211-23.

4.   Zhou L, McMahon C, Bhagat T, Alencar C, Yu Y, Fazzari M, et al. Reduced SMAD7 leads to overactivation of TGF-beta signaling in MDS that can be reversed by a specific inhibitor of TGF-beta receptor I kinase. Cancer research. 2011; **71**(3): 955-63.

5.   Gerstung M, Pellagatti A, Malcovati L, Giagounidis A, Porta MG, Jadersten M, et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. Nature communications. 2015; **6**: 5901.

 Appendices:

 None